Voluntary measures are an increasing part of the environmental policy portfolio in most of the world. The economics literature on voluntary measures has found mixed evidence of their effect on environmental outcomes, controlling for the behavior of those who are not affiliated with the measure (non-partners) and trends leading up to the policy initiation. Traditionally, a voluntary measure would be labeled a success if the measure's partners statistically improved their environmental outcome compared to non-partners, once the measure was initiated. However, that type of evaluation assumes that voluntary measures provide the treatment exclusively to partners. In practice, for many reasons, the treatment provided to partners may also spillover to affect the behavior of non-partners. In voluntary programs, information on the reuse of an input can spread from partner firms to non-partner firms. If the researcher does not account for these treatment spillovers in the evaluation, it is possible for successful measures to be found unsuccessful.

It is argued below that an alternative interpretation of a successful voluntary measure is needed, especially for those that involve treatment spillovers. Conditions under which a voluntary measure with treatment spillovers would be considered successful are discussed and then tested using the Coal Combustion Products Partnership (C2P2). The purpose of this voluntary program in the U.S. is to increase the reuse, as opposed to disposal, of coal combustion products (CCP). A difference-in-difference estimator finds that C2P2 partners are not statistically different from non-partners in their reuse rates, though the total reuse of coal combustion products has statistically increased over time. Further, *non*-partners located in states with more C2P2 partners statistically increase their reuse rate more than those in states with few C2P2 partners. The evidence is consistent with conditions needed for a program to be successful in reducing overall the disposal of CCP by partners and non-partners.

1. Background

The use of voluntary measures to improve environmental outcomes is common throughout the world, whether they be a voluntary program within one country or a voluntary agreement between countries. The literature suggests many reasons that firms or countries may join a voluntary program or agreement. They may join to improve their reputation with consumers or with voters (Khanna et al 1998; Arora & Cason, 1996) or to generate goodwill with the regulator or with other countries (Dawson & Segerson, 2008; Barrett 1994).

The evidence is mixed regarding the effectiveness of voluntary programs in the economics literature, although more often the evidence points to a lack of effectiveness. Theoretical analyses of voluntary program can be found in Lyon and Maxwell (2003) and Segerson and Miceli (1998).[1] Evaluations of numerous voluntary programs throughout the world can be found in Morgenstern and Pizer (2007), with most finding only a small, if any, effect on environmental outcomes. The U.S. has initiated a number of voluntary programs, beginning with the 33/50 program in 1991.[2] Khanna and Damon (1999), Innes and Sam (2008), and Artimura et al (2007) find improved environmental outcomes for the programs that they study (33/50, 33/50, and ISO 14001, respectively). Gamper-Rabindran (2006), Vidovic and Khanna (2006), and Brouhle et al (2008) find a lack of improvement in environmental outcomes for the programs they study (33/50, 33/50, and Strategic Goals Program for Metal Finishers, respectively).

The issue of whether a voluntary program induces treatment spillovers to non-partners is important for policy as well as for academic purposes. Voluntary programs are increasingly coming under scrutiny to show that they are the cause of improved environmental outcomes. The US Office of Management and Budget has and uses its authority over most voluntary programs (as well mandatory regulations) to ensure that public funds are being allocated efficiently. The US EPA Office of Inspector General has undertaken a number of analyses of voluntary programs in an attempt to encourage improvements in their operation. If voluntary program treatment spillovers are being ignored, oversight offices may find a lack of success in programs that are in fact successful and they may close those programs or curtail their funds and activities.

A similar pattern is revealed for evaluations of (voluntary) international environmental agreements. Finus and Tjotta (2003) and Murdoch and Sandler (1997) find that abatement targets for the Oslo and Montreal Protocols, respectively, were more in line with Nash equilibrium than with socially optimal targets. Bratberg et al (2005) and Swanson and Mason (2003) show that emissions from countries that signed the Sofia and Montreal Protocols, respectively, would have been larger in the absence of the protocols.

Recently, a number of studies have attempted to determine whether voluntary measures lead to increased innovation or changes in management of environmental systems. Carrion-Flores *et. al.* (2006) use patent application data to determine whether partners in the 33/50 program increased their innovative activity relative to non-partners.[3] Arimura *et. al* (2009) estimate whether firms that signed the ISO 14001 were more likely to require their suppliers to initiate an environmental management system. While the effects analyzed in the above studies are called spillovers, they primarily pertain to the behavior of partners. Thus they do not alter the evaluation of the program in the same way as

---

[1] This list of voluntary program evaluations is not exhaustive; it is meant only to give a general outline of the literature.

[2] A good background on voluntary programs in the U.S. can be found in Brouhle et al (2005).

[3] Dekker et al. (2009) run a similar analysis for the Helsinki and Oslo Protocols.

hypothesized in this analysis. The spillover discussed here pertains to the behavior of non-partners , and so alters the treatment effect being tested.

The traditional economic evaluation method for voluntary programs or agreements, as described in Khanna and Damon (1999) and Bratberg et al (2005), has labeled a program successful if those who are partners have statistically better environmental outcomes than those who are not. Lyon and Maxwell (2007) lay out a theory arguing that a different way to evaluate voluntary programs may be necessary for measures whose purpose is likely to cause information transfers (treatment spillovers). They argue that treatment spillovers may occur for many reasons. First, it is in the regulator's interest to have information disseminated as widely as possible if it will improve environmental performance. Second, information provided by a voluntary program may easily diffuse across an industry, making it difficult statistically to find a differential impact of the program on partners. The rate of diffusion will be higher when the information available through a voluntary program does not alter the competitive position of firms.

The traditional evaluation method is appropriate when it is expected that the treatment is only affecting partners and will not affect the behavior of non-partners. This interpretation would also be acceptable for a voluntary measure with weak treatment spillovers, so that the partners take advantage of the treatment to a larger extent than non-partners. However, an alternative interpretation for a successful voluntary program with treatment spillovers would be if two conditions are satisfied:

Condition 1: Environmental outcomes improve controlling for other factors and pre-measure trends.

Condition 2: Evidence shows that the treatment spillovers are affecting non-partners' behavior in a manner that improves their environmental outcome again controlling for other factors and pre-measure trends.

The first condition ensures that the voluntary measure is affecting behavior in a manner consistent with the goals of the program. The second condition ensures that the improved environmental outcome observed by non-partners would not have occurred in absence of the program. In essence, this means that it was the voluntary program that affected the environmental outcome of non-partners (and not other, non-program factors).[4]

---

[4] For a measure to affect environmental outcomes, only condition 2 needs to hold. However, this would imply that the partner's behavior is unaffected by the measure, while non-partner's

The above conditions are used to guide an evaluation of C2P2, a voluntary program to encourage the reuse of CCP that is housed in the Environmental Protection Agency's (EPA) Office of Solid Waste. It began in 2001 as an initiative and became a full program in 2003. C2P2 is part of the EPA's Resource Conservation Challenge, an attempt to encourage all members of an industry to achieve environmental outcomes similar to those of its cleanest member.[5] Currently, C2P2 has over 150 partners, including a number of trade associations, universities, federal agencies and private companies. The process of becoming a partner involves submitting a postcard with contact information to the EPA. The main benefit of joining C2P2 is the potential for increased CCP sales. Other benefits that are exclusive to partners are the ability to submit C2P2 award applications and to use the C2P2 logo. Other potential benefits of C2P2 are all available publicly to anyone, thus they are not a benefit of partnership, such as case studies or learning about past regulatory decisions.

Generally, the supply of CCP is from coal-fired power plants while the demand is from firms producing cement, aggregate, gypsum or other materials. The program accepts entities interested in CCP reuse, whether they are on the supply or demand side of a CCP reuse transaction. CCP are residuals from the coal combustion process such as fly ash, bottom ash and flue gas desfulurization wastes. C2P2 encourages reuse of CCP through educational workshops, case studies, facilitating research, and providing information on their uses and past regulatory decisions. Uses or re-uses of CCP include concrete, cement, drywall, asphalt, snow/ice control, and fill. The economic argument for C2P2 is that transactions cost inhibit this market from fully functioning in the absence of the program. These costs are generally thought to be search costs. For example, cement plants have to find nearby power plants and determine the regulatory liabilities in CCP reuse.

The goal of C2P2 is to increase the reuse ratio (reuse divided by total generation) of all CCP to 50% by 2011. The American Coal Ash Association (ACAA), whose mission is to encourage proper management and use of CCP, surveys power plants to collect data on production and use of CCP. The C2P2 program uses these data to track progress towards the goal of 50% reuse.[6] According to the ACAA and the Department of Energy (DOE) (2006), around 120 million tons of CCP are generated each year, making it one of the largest non-hazardous, non-municipal waste flows. Fly ash accounts for a little over half

---

behavior is affected. In this case, the purpose of the measure from a policy perspective is not being fulfilled.

[5] C2P2 also has the support of Power Partners, an electric utility partnership with the Department of Energy.

[6] However these data are not publicly-available at the plant level thus are not used in this analysis. The data used in this analysis have similar numbers for CCP, as is discussed below.

(55%) of the total CCP generated, with bottom ash accounting for 15%, and flue gas desulfurization material around 29%. CCP were initially exempted from the Resource Conservation and Recovery Act (RCRA), while the EPA studied whether they should be classified as hazardous. In 1993, the EPA determined that coal combustion products do not need to be regulated under RCRA. The existence of other federal and state programs dealing with solid wastes was listed as one of the reasons.

The direct goal of C2P2 is to increase the amount of CCP reused, but some reuses have additional environmental benefits. The largest category of reuse is fly ash as an input to concrete or cement products. An additional environmental benefit is that adding fly ash to concrete or cement production reduces the energy intensity and greenhouse gas emissions of the production process. The reuse of flue gas desulfurization waste reduces the energy intensity of the production of wallboard.

The C2P2 program fits the style of program that Lyon and Maxwell (2007) argue is likely to have program treatment spillovers. First, a large amount of information is available to partners and non-partners on the C2P2 webpage concerning reuse of CCP, such as past regulatory decisions and case studies. Second, information disseminated by C2P2 is unlikely to affect the competitive position of power plants because CCP disposal is a small fraction of power plants' costs, and because most power plants or utilities don't really compete with each other in the usual way, being regulated or geographically distinct. Third, C2P2 encompasses both suppliers and demanders of CCP in the program. A scenario can be imagined where a potential demander of CCP learns more about them through C2P2 and contacts a local power plant that is not a member of C2P2 to discuss purchasing CCP. The resulting increase in the reuse of CCP would be attributed to a non-partner power plant in this analysis, though the impetus for the reuse came from C2P2 information. It is this third method of treatment spillover that this analysis will exploit to test whether C2P2 spillovers are improving non-partners' reuse of CCP.

2. Data

Data for this analysis come from the Energy Information Administration (EIA) and the United States Geological Survey (USGS). Voluntary programs are generally difficult to evaluate due to the lack of data available before the program started and for non-partners once the program is in effect. However, the EIA has been collecting information on CCP for many years as part of its Form 767: Annual Steam-Electric Plant Operation and Design Data. Observations used here are from the years 1996-2005. In 2001, the EIA began collecting information from smaller boilers (<25 MW) in Form 767. These smaller boilers are removed from the sample due to their lack of information before the C2P2 program started.

As the dependent variable, this analysis employs the ratio of fly ash reused to total fly ash generated.[7] As discussed above, fly ash is the largest CCP category and accounts for 55% of all CCP. The EIA asks plants to report total by-products generated, the amount landfilled on-site (both wet and dry), the amount landfilled offsite, the amount used or stored on-site, and the amount sold.[8] The fly ash reuse ratio is calculated as the amount of fly ash sold divided by the total by-product fly ash generated.

The C2P2 webpage lists all partners, though it does not list the date at which each firm became a partner. However, the Utilities Solid Waste Activities Group (a trade association) website lists the firms that initially committed to C2P2.[9] A list of these utilities is given in Table 1. For this analysis, the utilities listed in Table 1 are considered partners. The utilities listed on the C2P2 webpage but not in Table 1 are called "late partners" and are excluded from the econometric analysis due to the lack of information on their year of entrance to C2P2. The balance of utilities is considered non-partners. It is assumed that if a utility is a partner (non-partner), then all the plants it owns are partners (non-partners).[10]

Figure 1 gives the reuse ratio for fly ash, by C2P2 partner designation, over the sample years. ACAA information on reuse shows that the average reuse ratio for all plants is 45% in the early 2000s. The average reuse ratios found with the sample used here from the EIA-767 data is higher (47%). The figure reveals that initial partners were generally re-using less fly ash than non-partners and late partners, with late partners having the highest reuse ratio. This pattern suggests that the initial partners may have initiated C2P2 due to their difficulty in re-using CCP while the late partners were originally the industry leaders in re-using CCP. The pattern of partner timing choice with C2P2 has been observed in other programs (Delmas and Montes, 2007).

---

[7] Information on other CCP are not consistent enough to analyze.

[8] Most CCP are either sold or landfilled onsite.

[9] http://www.uswag.org/c2p2.htm, last accessed 3/10/08

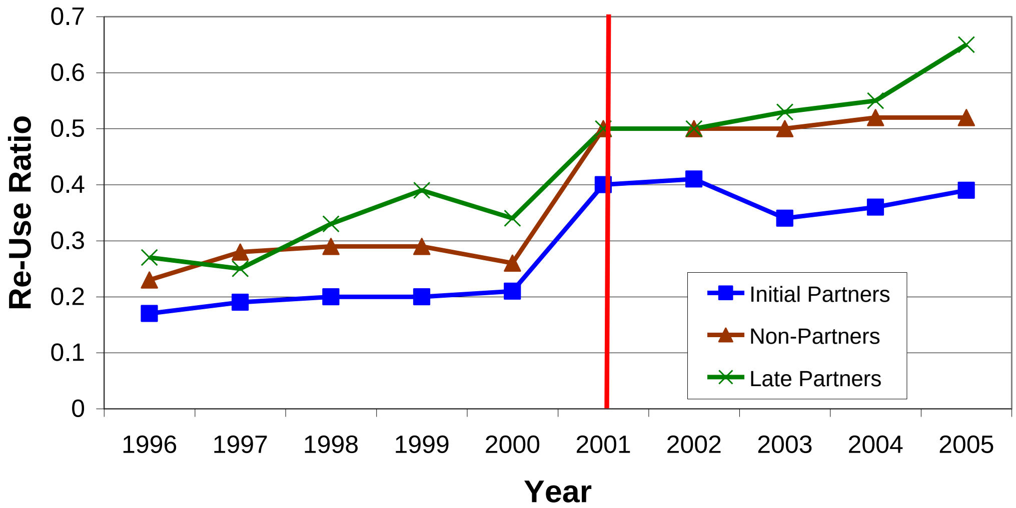
[10] Based on personal communication with John Sager, the lead for the C2P2 program

Table 1: Initial C2P2 Partners
Alliant Energy
Ameren Corporation
American Electric Power Company
Constellation Energy Group
Consumers Energy
Duke Energy
FirstEnergy
Indianapolis Power & Light Company
LG&E Energy Corporation
Mirant Corporation
Montana-Dakota Utilities Company
Progress Energy
Public Service Company of New Hampshire
Public Service Enterprise Group
Reliant Energy
Southern Company
Tennessee Valley Authority
Tri-State Generation & Transmission
Xcel Energy

The explanatory variables in this analysis come from the EIA-767 and the USGS. Explanatory variables from the EIA-767 data are the annual coal consumption (in 100,000 tons), ash content of the coal burned, fly ash reuse of other plants owned by the same utility, location of other plants owned by the same utility, and the presence of a selective catalytic reduction (SCR) technology at the plant. The ash content is the average ash content, in percent by weight, of the coal burned in the year. The average fly ash reuse for all plants owned by a utility except the one in question is calculated, and then estimated in lags, in order to control for potential learning over time by the utility itself. If the utility owns only one plant, then this variable is equal to zero. The number of plants owned by a utility in the same state as the one in question is calculated to control for the potential for excess demand at one plant being given to other plants related by ownership. An SCR is a nitrogen oxide (NOx) pollution control device that can lower the quality of the resulting fly ash.[11]

[11] Mercury control through activated carbon injection can also affect the quality of the resulting fly ash. However, none of the plants during our sample years used activated carbon injection.

Figure 1: Fly Ash Reuse Ratios over Time



Information was also gathered from the USGS Mineral Yearbook. Fly ash can be a substitute for cement and crushed stones or aggregates. The average value of cement per state in dollars per metric tons is taken from Table 11 of the USGS Cement Minerals Yearbook for the years 1996-2005. The average value of crushed stone in dollars per metric ton per state is taken from Table 4 of the USGS Crushed Stone Minerals Yearbook for the years 1996-2005. During the sample time period, the cement industry was operating close to capacity. Fly ash reuse and cement imports, two close substitutes for domestic cement, rose to meet the excess demand. Thus, the level of cement imports for each year is taken from the USGS Cement Mineral Yearbook to control for the effect of excess demand on the reuse of fly ash. Finally, nine regional dummy variables based on Census Division regions and ten annual dummy variables are constructed with the region and year in question taking the value of one and otherwise zero.

Variables specifically relating to the evaluation of the C2P2 program are based around the partner designation described above and suggestions given by Lyon and Maxwell (2007). The sample time period is split into three variables: pre-C2P2, early-C2P2, and late-C2P2. Each variable is equal to one during the corresponding years and is zero otherwise. The pre-C2P2 time period dummy is equal to one in the years 1996-2000, before C2P2 was formed. The early-C2P2 period dummy is equal to one in the years 2001-2002, and the late-C2P2 period dummy is equal to one in the years 2003-2005. The early-C2P2 and late-C2P2 variables are then interacted with the partner variable to determine whether partners' behavior is significantly different from non-partners' behavior. This structure only allows for temporal variation in the treatment (C2P2) effect.[12]

3. Analysis

Summary statistics for the variables used in this analysis can be found in Table 2 for all groups and by partner designation. The estimation sample includes 127 partner plants and 177 non-partner plants. The information presented in Figure 1 provides evidence of a pattern in the choice of partner designation and imply that partners seem to have a more difficult time initially re-using fly ash than non-partners. Table 3 confirms that the fly ash reuse ratios between the three groups are statistically different from each other using a t-test. Table 2 reveals other patterns, such as non-partners on average burning lower ash coal and facing a lower price for aggregates. The price of cement and coal burned per plant generally does not differ by partner designation. Given the pattern of the dependent and independent variables and the potential for selection bias in becoming a partner, this analysis will first generate a prediction as to whether a plant will participate in C2P2.

A probit analysis is undertaken to ensure that the choice of participation in C2P2 does not bias evaluation results. A prediction of participation is created using instruments that will not be in the evaluation estimation together with variables that will be in the evaluation estimation. The dependent variable in the participation analysis is the initial partner variable. The explanatory variables for the participation analysis include three variables that are also in the reuse analysis (though not in the same form): the average fly ash reuse ratio, the average price of cement, and the average price of aggregates for the years 1996-2000. Each is the averages of the variables across the years 1996-2000. It is expected that a higher fly ash reuse ratio for the years 1996-2000 would make it less likely a plant joins the program, in light of the information in Figure 1. Higher cement and aggregate prices would seem to have an ambiguous effect on the likelihood of a plant joining, since the higher price would encourage CCP demanders to seek out the power plant, while also bringing in more revenue to the plant if more CCP are reused.

[12] Based on personal communication with those running the C2P2 program, information on the year which late partners joined the program does not exist. Internet searches for this information from the utilities themselves were not successful.

Table 2: Summary Statistics

	Estimation Sample		Initial Partners		Non Partners		Late Partners	
	n=2242		n=949		n=1293		n=485	
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Fly Ash Reuse Ratio	0.31	0.35	0.26	0.33	0.36	0.37	0.40	0.35
Restructured Market	0.36	0.48	0.34	0.47	0.36	0.48	0.47	0.50
Utility Size (# of boilers)	8.80	12.00	12.80	15.40	6.03	7.59	7.50	7.63
NO_x SIP Call State	0.53	0.22	0.07	0.25	0.04	0.20	0.10	0.30
Re-Use Authorized	0.86	0.34	0.92	0.27	0.79	0.40	0.86	0.34
No CCP Permit Required	0.28	0.45	0.34	0.47	0.23	0.42	0.29	0.45
Solid Waste Investment 96-00 (100,000 $)	10.60	32.50	13.50	40.22	8.85	25.50	12.60	29.20
Total Coal (100,000 tons)	11.20	9.90	11.10	10.10	11.60	10.08	9.80	9.01
Ash Content (% by weight)	8.03	3.70	8.38	3.21	7.78	4.01	7.84	4.31
USWAG Member	0.43	0.49						
Utility Reuse	0.16	0.22	0.13	0.17	0.17	0.22	0.21	0.22
Number of Utility's Plants Near	2.45	2.23	3.33	2.40	1.71	1.81	2.57	2.08
Aggregate Price ($ per ton)	5.23	1.26	5.58	1.22	5.10	1.45	5.50	1.01
Cement Price ($ per ton)	75.30	5.03	74.30	4.40	76.01	5.32	74.30	6.45
Cement Imports (million tons/year)	23.70	5.24	23.68	5.24	23.78	5.18	23.95	5.12
SCR Installed	0.03	0.16	0.04	0.18	0.03	0.15	0.04	0.20

Note: Late Partners are not included in the Estimation Sample

Six exogenous variables are used to instrument for participation and are only used in the participation analysis. First is the utility size, measured by the number of boilers a utility owns, as given in the EIA-767 data. Most of the literature on voluntary programs finds that larger firms are more likely to join a voluntary program. Second is the average bottom ash reuse ratio for the years 1996-2000, given in the EIA-767 data. Similar to the expected effect of fly ash reuse rates over this time period, it is expected that plants with lower bottom ash reuse ratios are more likely to join C2P2. Third is the total amount invested in solid waste disposal at the plant for the years 1996-2000, as given in the EIA-767 data. It is presumed that plants with less invested in solid waste disposal would be more interested in re-using their CCP and thus more likely to join C2P2. Fourth is whether the state in which the plant is located has authorized CCP reuse in some form as given by the DOE National Energy Technology Lab. Fifth is whether the state in which the plant is located exempts CCP from solid waste permitting requirements (DOE, 2006, Table 20). The authorization of CCP reuse and exemption from solid waste permitting are steps that states take to encourage reuse, so these two variables are expected to increase the likelihood of joining C2P2.[13] Sixth and final is whether the state a plant is located in has restructured its electricity market, according to the EIA (2003). No expectation as to how the restructuring of electricity markets would affect the likelihood of joining C2P2 is

[13] If CCP is not an exemption the regulations surrounding the reuse are more complicated legally. This information comes from personal conversations with John Sager, the lead for the C2P2 program.

proposed. All of these dummy variables take the value of one to indicate plants located in states that meet the criteria listed above and is zero otherwise.

Table 3: Hypothesis Tests

Fly Ash Reuse Ratio T-Tests

	Partner	Non-Partner	Late Partner
Partner		6.19	8.11
Non-Partner	6.19		3.21
Late-Partner	8.11	3.21	

Null Hypothesis: Row Fly Ash Reuse = Column Fly Ash Reuse

A probit model is used to predict partner designation, which takes the following form:

$$\mathbf{P_i} = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1\mathbf{R}_i + \boldsymbol{\beta}_2\mathbf{S}_j + \boldsymbol{\beta}_3\mathbf{I}_i + \boldsymbol{\mu}_i \quad [1]$$

where P_i is a binomial variable equal to one if the plant was an initial partner and zero if it is not a partner (late partners are excluded from this analysis), R_i is the average reuse ratio for fly and bottom ash for the years 1996-2000, S_j is a vector of the average price of cement and aggregates for the years 1996-2000, I_i is the vector of six instruments described above, and μ_i is an error term.

The results from Equation 1 are then used as part of the evaluation of C2P2. The evaluation is carried out as the two conditions for a successful voluntary program with treatment spillovers, discussed above, specify. First, a difference-in-difference model is estimated to determine whether the environmental outcome of partners and non-partners has increased. Second, a test is undertaken for evidence consistent with treatment spillovers leading to improved environmental outcomes for non-partners. The assumption that non-partners act as a control, an unaffected group, is by definition relaxed. In its place, non-partner power plants who are located in states with few C2P2 CCP demand partners (non power plant C2P2 partners) act as the control.

To test the first condition, we hypothesize that the level of fly ash reuse by plants is a function of: the total coal burnt, the ash content of the coal, fly ash reuse of other plants owned by the same utility, location of other plants owned by the same utility, the price of cement and aggregates in the state, the level of cement imports, the presence of an SCR at the plant, plant specific effects and the information disseminated by C2P2. Information dissemination by C2P2 includes educating state and local agencies, conducting research on reuse applications, and discussing the benefits of CCP reuse to potential demanders. This information is likely to impact the reuse decisions of partners and non-partners. In order to determine whether the C2P2 program has led to increased reuse of coal combustion products a difference-in-difference fixed effects model is estimated. The evaluation model is given in Equation 2:

$$R_{it} = \alpha_i + \beta_4 X_{it} + \beta_5 S_{jt} + \beta_6 \hat{P}_i + \beta_7 T_{it} + \beta_8 T \hat{P}_{it} + \varepsilon_{it} \qquad [2]$$

Where R_{it} is the reuse ratio for fly ash for plant i at time t, X_{it} is a vector of plant related variables, S_{jt} is a vector of cement and aggregate variables by state, $\hat{P}_i$ is the predicted partner variable, T_{it} is a vector of C2P2 period dummies, $T\hat{P}_{it}$ is an interaction of C2P2 period and predicted partner variables (the difference-in-difference parameter), α_i is the fixed effects parameter, and ε_{it} is an error term. For this evaluation, the value of β_8 is the policy variable of interest. It will reveal whether predicted partner plants changed their reuse ratio after C2P2 went into affect relative to plants that are not predicted to enter the program.

Other specifications estimated in addition to fixed effects are a random-effects, dynamic panel, panel-corrected first order auto-regressive standard errors specification, and random effects ordered probit, though not all are discussed in the text.[14] The dynamic panel specification corrects for potential learning by doing within the plant. The panel-corrected first order auto-regressive standard errors specification corrects for potential persistency of the error term within panel. The random effects ordered probit ensures that the dependent variable is not predicted to be outside of the zero-one range.

Equation 2 is estimated separately using the predicted partner variables and actual partner variables. In both cases, the standard errors are clustered by utility to control for potential within-utility correlations over time, given that the data are a panel of plants owned by different utilities and that utilities may own more than one plant. When the predicted partner variables are used, the standard errors are bootstrapped following Vidovic & Khanna (2006), due to the use of estimated independent variable in the specification.

In order to test for evidence of treatment spillovers to non-partners, information is matched between location of C2P2 CCP demand (non-power plant) partners and the location of non-partner power plants. The scenario discussed above, an example of treatment spillovers, is where a potential demander of CCP learns more about them through C2P2, contacts a local power plant that is not a member of C2P2, and begins purchasing their CCP. If this scenario occurs, it is expected that non-partner power plants in states with many C2P2 CCP demand

[14] The estimation results are not shown though they are available by request. Results for the dynamic panel come from the xtabond2 command from Rodman (2009).

partners would increase their reuse rates compared to non-partner power plants in states that have few C2P2 CCP demand partners.

This hypothesis is tested by separating states into two categories, those with low C2P2 CCP demand partners, and high CCP demand partners. The states Arizona, Colorado, Kentucky, Michigan, Missouri, North Dakota, Ohio, Pennsylvania, Texas, and Wisconsin are high CCP demand partner states. They have a large number of C2P2 CCP demanders as partners and most have a state university center interested in CCP issues. The low C2P2 CCP demand partner states are: Delaware, Illinois, Indiana, Iowa, Kansas, Louisiana, Massachusetts, Minnesota, Mississippi, Nebraska, Nevada, New Mexico, New York, Oklahoma, South Carolina, South Dakota, Utah, Virginia, West Virginia and Wyoming. A low C2P2 CCP demand partner dummy variable is set to one in those states, and it is set to zero otherwise. Equation 2 is re-run for non-partners exclusively, with the high C2P2 CCP demand partner dummy interacted with the C2P2 time period dummies. This interaction will reveal if non-partner power plants that are more likely to receive program information have different reuse rates once C2P2 was in effect from non-partner power plants that are less likely to receive program information.[15]

4. Results

The results of the estimation of equation 1 are listed in Table 4. A number of the instruments are statistically significant. If the state has authorized reuse of CCP or an exemption for CCP from solid waste permitting requirements, then the plant is more likely to join C2P2, as expected. However, the amount invested in solid waste disposal has no statistical impact on the decision to join the program, perhaps because it is a sunk cost. Past reuse rates for fly ash have a statistically significant impact on the decision to join C2P2, consistent with the pattern in Figure 1 and Table 3. A higher price for aggregates significantly increases the probability that the plant is a partner, but higher cement prices have the opposite effect.

---

[15] We would like to remind the reader that the data are aggregated at the plant level, making any attempt to determine characteristics of specific fly ash transactions impossible.

Table 4: Participation Regression Results

Dependent Variable:	Initial Partner Dummy	
Model:	Probit	
Explanatory Variable	Coefficient	Stnd. Error
Avg Fly Ash Reuse Ratio 1996-2000	-0.51*	0.28
Avg Bottom Ash Reuse Ratio 1996-2000	0.41	0.24
Avg Aggregates Price 1996-2000	0.34***	0.08
Avg Cement Price 1996-2000	-0.02	0.02
Utility Size	0.01	0.01
Restructured Market	-0.48**	0.18
Total Solid Waste Disposal Investment 1996-2000	0.03	0.02
No CCP Permit Required	0.45**	0.21
Re-use Guidelines Set	0.81***	0.26
N	303	
R2	0.11	

*, **, *** indicate 10%, 5% and 1% statistical significance, respectively

These results are then used to predict partner designation in the C2P2 program. This predicted partner designation variable is used to estimate equation 2. Table 5 lists the results for a fixed and random effects specification using the predicted partner and a random effects specification, with actual partner designation. All three estimations reveal the same pattern; the rate of fly ash reuse increased after C2P2 went into effect for both partners and non-partners. These increases occur in both the early-C2P2 and the late-C2P2 variables, as they are positive and statistically significant. However, when these variables are interacted with the predicted initial partner or the actual initial partner variable, the coefficients are not statistically different from zero. This implies that plants that were initial partners or those with a higher predicted probability of being a partner do not increase their reuse rate relative to those that are not partners or those with a lower probability of being a partner. Under the traditional interpretation of a voluntary program evaluation, C2P2 would look like it has no effect. However, if evidence of program spillovers could be found, the interpretation would be different.

Table 5: Evaluation Regression Results

Time Period : 1996-2005						
Dependent Variable:	Fly Ash Reuse Ratio		Fly Ash Reuse Ratio		Fly Ash Reuse Ratio	
Model:	Fixed Effects		Random Effects		Random Effects	
Explanatory Variable	Coefficient	Stnd. Error	Coefficient	Stnd. Error	Coefficient	Stnd. Error
Predicted Partner			-0.18**	0.08		
Actual Parner					-0.02	0.04
Early-C2P2 (2001-2002)	0.06**	0.02	0.08**	0.03	0.06**	0.03
Late-C2P2 (2003-2005)	0.06**	0.02	0.07**	0.02	0.09**	0.03
Early-C2P2* Predicted Partner	0.03	0.02	0.03	0.02		
Late C2P2* Predicted Partner	-0.01	0.02	-0.01	0.02		
Early-C2P2* Actual Partner					-0.01	0.04
Late C2P2* Actual Partner					-0.05	0.04
Number of Utility's Plants Near	0.01*	0	0.01	0.01	-0.01	0.01
Post-C2P2* Utility's Plants Near	0.01	0.01	0.01	0.01	0.01	0.01
Lag Utility Reuse	0.08**	0.04	0.11***	0.03	0.11**	0.05
Aggregates Price	-0.01	0.02	0.06***	0.02	0.03**	0.01
Cement Price	0.01	0.01	-0.01	0.01	-0.01	0.01
Cement Imports	0.01	0.01	0.01	1	0.01	0.01
SCR Installed	0.01	0.02	0.01	0.01	0.02	0.01
Avearge Ash Content	-0.01***	0	-0.02***	0.01	-0.02***	0.01
Total Coal	0.01**	0.00	0.01***	0.00	0.01**	0.00
N	1883		1883		1883	
Plants	303		303		303	
R2	0.13		0.3		0.3	

Region & Year Dummies Not Shown for Brevity
Errors Clustered by Utility for all specifications and Bootstrapped for the first two specifications
*, **, *** indicate 10%, 5% and 1% statistical significance, respectively

The model shown in Table 5 controls for potential learning-by-doing within the utility through the lag utility reuse variable, but not by the plant. Learning-by-doing could also occur within the plant, which could be controlled by using a lagged dependent variable as an independent variable, but such a step would require the use of a dynamic panel estimator as outlined in Arellano and Bond (1991). This method requires the use of exogenous or predetermined variables to help identify the effects of the lagged dependent variable on the contemporary dependent variable. Results for a dynamic panel model generally reject the identification variables used (i.e., Sargan Test) or they find that the lagged dependent variable is not statistically significant. The coefficients on the C2P2 time period variables from a dynamic panel specification are not as statistically consistent as those from the specifications shown. Some specifications show no change in environmental outcomes after C2P2 went into effect, while others show a statistically significant improvement.

Given the potential for treatment spillovers with this program, a comparison is made of non-partner power plants close to many demand partners to those close to few demand partners. The results for this program spillovers analysis are given in Table 6. Consistent with Table 5, non-partners increased their reuse rate after C2P2 went into effect, given by the positive and significant coefficient of the early-C2P2 and the late-C2P2 variables. However, plants in states with low C2P2 CCP demand partners have a statistically significantly smaller increase in reuse rates. This outcome is consistent with a treatment spillover from C2P2 CCP demand partners to non-partner power plants.

If C2P2's treatment spills-over from partners to non-partners, then a traditional difference-in-difference analysis that compares the change for partners to the change for non-partners may be biased toward finding no statistically significant effect, even if the program does effect both groups. This analysis finds that fly ash reuse has statistically significantly increased for partners and non-partners and, further, non-partners with many C2P2 CCP demander partners located in the state increased their reuse rate at a statistically larger rate than those non-partners located in states with low C2P2 CCP demander partners. This evidence is consistent with the conditions laid out for a successful voluntary program.

Table 6: Spillover Test Results

Time Period : 1996-2005				
Dependent Variable:	Fly Ash Reuse Ratio		Fly Ash Reuse Ratio	
Model:	Fixed Effects		Random Effects	
Explanatory Variable	Coefficient	Stnd. Error	Coefficient	Stnd. Error
Moderate/Low C2P2 Partners Nearby			0.34***	0.08
Early-C2P2 (2001-2002)	0.10**	0.04	0.11**	0.04
Late-C2P2 (2003-2005)	0.14**	0.05	0.15**	0.05
Early-C2P2* Low C2P2 Partner	-0.09*	0.05	-0.09*	0.05
Late-C2P2* Low C2P2 Partner	-0.14**	0.06	-0.14**	0.06
Number of Utility's Plants Near	-0.01	0.01	-0.01	0.01
Post-C2P2* Utility's Plants Near	0.01	0.02	0.01	0.02
Lag Utility Reuse	0.07	0.06	0.04	0.06
Aggregates Price	-0.01	0.02	-0.01	0.02
Cement Price	-0.01	0.01	-0.01*	0.00
Cement Imports	0.01	0.01	0.01	0.01
SCR Installed	0.04	0.03	0.04	0.04
Avearge Ash Content	-0.01**	0	-0.02**	0.01
Total Coal	-0.01	0.01	0.01**	0.00
N	1088		1088	
Plants	176		176	
R2	0.2		0.48	

State & Year Dummies Not Shown for Brevity

Errors Clustered by Utility

*, **, *** indicate 10%, 5% and 1% statistical significance, respectively

5. Conclusion

Economic analyses of voluntary programs have found mixed evidence that they improve the environmental performance of firms in the program compared to those not in the program. Lyon and Maxwell (2007) argue that this result may occur because the programs have treatment spillover effects that affect non-partners as well as partners. The traditional interpretation of a voluntary program evaluation states that partners must have a better environmental outcome than non-partners for the program to be successful. If the program induces treatment spillovers, then this traditional view is unlikely to be found (and indeed may lead to incorrect interpretations). In the case of treatment spillovers, an evaluation should find that a) both partners and non-partners improve their environmental outcome and b) evidence that treatment spillovers are affecting non-partners behavior, controlling for pre-program trends and other determinants of the outcome.

The C2P2 program is likely to have treatment spillovers due to the fact that the program includes both suppliers and demanders of CCP. A scenario where a C2P2 CCP demand partner would learn of CCP benefits from C2P2, and then transact with a non-C2P2 partner for supply is quite possible. However, the increased reuse of CCP in this scenario is attributable to the C2P2 program even though it looks like a non-partner's reuse has increased. An evaluation of the C2P2 program is performed with a difference-in-difference estimator to determine whether C2P2 partners improved their reuse of CCP at a statistically significantly different rate. Results generally find that C2P2 partners are no different from non-partners in their reuse rates, though the total reuse of CCP has statistically increased since the C2P2 program went into effect. If this program were evaluated using the traditional conditions of the literature, the results would be consistent with an unsuccessful program. However, evidence also points to a larger increase in reuse for non-partner power plants located near many other demand partners, relative to those that are not near other C2P2 CCP demand partners. The design of the program along with the empirical results are consistent with the conditions given for a program with treatment spillovers to be successful.

References

American Coal Ash Association (Various Years) "ACAA Coal Combustion Products Survey". www.acaa-usa.org.

Arellano, Manuel and Stephen Bond. 1991. Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *The Review of Economic Studies* 58 (2), 277-297.

Arimura, Toshi, Nicole Darnall, and Hajime Katayama. 2009. Is ISO 14001 a Gateway to More Advanced Voluntary Action? A Case for Green Supply Chain Management. RFF Discussion Paper 09-05.

Arimura, Toshi, Akira Hibiki, and Hajime Katayama. 2008. Is a Voluntary Approach an Effective Environmental Policy Instrument? A Case for Environmental Management Systems. *Journal of Environmental Economics and Management* 55(3), 281-295.

Arora, Seema & Timothy Cason. 1996. Why Do Firms Volunteer to Exceed Environmental Regulations? Understanding Participation in EPA's 33/50 Program. *Land Economics* 72(4), 413-432.

Barrett, Scott. 1994. Self-enforcing International Environmental Agreements. *Oxford Economic Papers* 46(3), 878–884.

Bratberg, Espen & Sigve Tjotta & Torgeir Oines. 2005. Do Voluntary International Environmental Agreements Work? *Journal of Environmental Economics and Management* 50(3), 583-597

Brouhle, Keith, Charles Griffiths and Ann Wolverton. 2005. "The Use of Voluntary Approaches for Environmental Policymaking in the U.S." in *The Handbook of Environmental Voluntary Agreements*, E. Croci, ed. Kluwer Academic Publisher

Brouhle, Keith, Charles Griffiths and Ann Wolverton. 2009. Evaluating the Role of EPA Policy Levers: An Examination of a Voluntary Program and Regulatory threat in the metal finishing industry. *Journal of Environmental Economics and Management* 57(2), 166-181.

Carrion-Flores, Carmen. Robert Innes, and Abdoul Sam. 2006. Do Voluntary Programs Spur Innovation in Environmental Technology? American Agricultural Economics Association Annual Meeting.

Dawson, Na Li & Kathy Segerson. 2008. Voluntary Agreements with Industries: Participation Incentives and with Industry-Wide Targets. *Land Economics* 84 (1), 97-114.

Dekker, Thijs, Herman Vollebergh, Frans de Vries, & Cees Withagen. 2009. How International Environmental Agreements Trigger Knowledge Transfers: Evidence from SO2 Protocols. 2009 European Association of Environmental and Resource Economics Annual Meeting.

Delmas, Magali and Maria Montes. 2007. Voluntary Agreements to Improve Environmental Quality: Are Late Joiners the Free Riders? UCSB ICER Working Paper 07.

Finus, Michael & Sigve Tjotta. 2003. The Oslo Protocol on Sulfur Reduction: The Great Leap Forward? *Journal of Public Economics* 87(9-10), 2031-2048.

Gamper-Rabindran, Shanti. 2006. Did the EPA's Voluntary Industrial Toxics Program Reduce Emissions? A GIS Analysis of Distributional Impacts and by Media Analysis of Substitution. *Journal of Environmental Economics and Management* 52(1), 391–410.

Innes, Robert and Abdoul Sam. 2008. Voluntary Pollution Reductions and the Enforcement of Environmental Law: An Empirical Study of the 33/50 Program. *Journal of Law and Economics* 51(2), 271-296.

Khanna, Madhu and Lisa Damon. 1999. EPA's Voluntary 33/50 Program: Impact on Toxic Releases and Economic Performance of Firms. *Journal of Environmental Economics and Management* 37(1), 1-25

Lyon, Thomas and John Maxwell. 2003. Self-Regulation, Taxation and Public Voluntary Environmental Agreements. *Journal of Public Economics* 87(7-8), 1453-1486.

Lyon, Thomas and John Maxwell. 2007. Public Voluntary Programs Reconsidered. *Policy Studies Journal* 35(4), 723-750.

Morgenstern, Richard and William Pizer, editors. 2007. *Reality Check: The Nature and Performance of Voluntary Environmental Programs in the United States, Europe and Japan*, RFF Press.

Murdoch, James and Todd Sandler. 1997. The Voluntary Provision of a Pure Public Good: the Case of Reduced CFC Emissions and the Montreal Protocol. *Journal of Public Economics* 63(2), 331-349.

Roodman, David. 2009. How to do Xtabond2: An Introduction to Difference and System GMM in Stata. *Stata Journal* 9(1), 86-136.

Sam, Abdoul and Robert Innes. 2008. Voluntary Pollution Reductions and the Enforcement of Environmental Law: an Empirical Study of the 33/50 Program. *Journal of Law and Economics* 51 (2), 271-296

Segerson, Kathleen and Thomas Miceli. 1998. Voluntary Environmetnal Agreements: Good or Bad News for Environmental Protection? *Journal of Environmental Economics and Management* 36(2), 109-130.

Swanson, Timothy and Roger Mason. 2003. The Impact of International Environmental Agreements: The Case of the Montreal Protocol. In: Marsiliani, L., M. Rauscher und C. Withagen (Hrsg.), *Environmental Policy in an International Perspective*, Kluwer, Dordrecht, Holland, 51-80.

U.S. Energy Information Administration (Various Years) Form 767: Annual Steam-Electric Plant Operation and Design Data

U.S. Energy Information Administration (2003) "Status of State Electric Industry Restructuring Activity."

U.S. Department of Energy (2006) "Coal Combustion Waste Management at Landfills and Surface Impoundments, 1994-2004." DOE/PI-0004

U.S. Geological Service. (Various Years) "Coal Combustion Products Commodity Report."

Vidovic, Martina and Neha Khanna. 2006. Can Voluntary Pollution Prevention Programs Fulfill their Promises? Further evidence from the EPA's 33/50 Program. *Journal of Environmental Economics and Management* 53(2), 180-195.